How the Preference Reversal Phenomenon Appears and Disappears.

CHRISTOPHE SCHMELTZER[1*], JEAN-PAUL CAVERNI[1] and MASSIMO WARGLIEN[2]

[1]*Cognitive Psychology Laboratory, Aix-Marseille I University and CNRS, Marseille, France*

[2] *Cognitive Science Lab, Rovereto* and *Department of Business, Economics and Management, Ca'Foscari University, Venezia, Italy*

* Correspondence to: Christophe Schmeltzer, Université d'Aix-Marseille I, Laboratoire de Psychologie Cognitive, Case 66, 3 Place Victor Hugo, F-13331 Marseille Cedex 3, France. Tel: +33491108517. Fax: +33442965293. E-mail: christophe.schmeltzer@up.univ-mrs.fr

ABSTRACT

The preference reversal phenomenon is a systematic change in the preference order between options through different situations (e.g. choice vs. judgment). The present study focuses on procedures used to elicit preferences, according to an evaluability hypothesis. Two experiments compared Joint vs. Separate evaluations and Explicit vs. No explicit Joint evaluations. Subjects had to express preferences between high-variance gambles (HVG) and low-variance gambles (LVG) either by choosing one gamble to participate in a lottery or by assigning them a minimum selling price.

Results showed that HVG are preferred in both choice and selling conditions when gambles were evaluated separately with numerical response scales, although LVG are preferred in both choice and selling conditions when gambles were evaluated by pair with binary response scales.

Such results sustain the evaluability hypothesis. They suggest that preferences depend on the way subjects are allowed or not to effectively compare the options between which they are asked to choose or to judge, independently of the nature of the scale (i.e. attractiveness vs. minimum selling price) they are asked to use.

KEY WORDS: preference reversal; evaluability theory; anchoring process

INTRODUCTION

The preference reversal (PR) phenomenon has been on stage for more than 30 years now, since the initial work by Lichtenstein and Slovic (1971). It can be broadly defined as a change in the preference order between options when different procedures are used to elicit such preferences. The best-known example involves lottery gambles in choice vs. judgment procedures: subjects are presented with two different gambles of equal (or nearly equal) expected value. Although they usually prefer low-variance gambles (called "P-bet" in the PR literature) with higher probability of winning small outcomes when choosing which of the two gambles is more attractive, they prefer high-variance gambles (called "$-bet" in the PR literature) with lower probability of winning high outcomes when evaluating the minimum selling price of each gamble. We will refer to this baseline result as the “classic” PR.

This phenomenon is just one of the four possible types of individual responses when a same pair of low/high-variance gambles is evaluated with different elicitation procedures. Exhibit 1 represents these four possibilities: when low-variance gambles (LVG) are preferred in both procedures, subjects exhibit consistently risk-averse behavior; consistently risk-seeking behavior is displayed when high-variance gambles (HVG) are preferred in both procedures; when LVG are preferred in choice procedure while HVG are preferred in judgment procedure, subjects show the classic PR; the opposite behavior is rarely observed. Of course, in any experiment, all four response types are observed, but the modal one (usually between 40% and 50% of the responses) is the classic PR.

[insert Exhibit 1 around here]

Over time, this phenomenon has demonstrated its robustness, resisting to many experimental attacks (cf. Grether and Plott, 1979; Pommerehne *et al*., 1982; Reilly, 1982)[1] and has been replicated in a growing set of decision tasks. This has not come without paying the price of increasing complexity: multiple factors generating PR have emerged, altogether with different theoretical explanations trying to account for the phenomenon. For extensive reviews, see Slovic and Lichtenstein (1983), Tversky *et al*. (1990), and Camerer (1995). For Caverni (1996), the understanding of the processes underlying PR might be improved by taking into account all the features in which the two tasks (judgment vs. choice) differ. Pointing out that the response for choice is a comparative one (binary response scale), while the response for the pricing is an absolute one (numerical response scale), he suggested to ask subjects to make absolute choices and do comparative pricing. Our research draws its inspiration from this suggestion.

If PR can be caused by many different factors, interactions between them should matter. However, such interactions are by now under-investigated, Goldstein and Einhorn (1987) still being the most relevant exception. Goldstein and Einhorn proposed to distinguish two dimensions of the procedures involved in these classical decision making tasks: the "*response method*", ("[…] what subjects have been asked to do, i.e., choose or

[1] By using a procedure (simultaneous tasks) which forced subjects to debias their judgements and made explicit their discrepancies, Ordóñez *et al*. (1995) observed that the phenomenon was significantly reduced. In this paper, we will refer only to the classical conditions (sequential tasks).

judge", p. 237) and the "*worth scale*" ("[…] the scale the subjects have been asked to do it with, i.e., attractiveness or minimum selling price", idem). This distinction leads them to define different types of PR, according to different combinations of the response method and the worth scale.

Goldstein and Einhorn's distinction has been recently developed by Hsee *et al*. (1999) to clarify the role of the response method. Hsee *et al*. suggest that a basic source of differentiation between response methods is due to the degree of comparativeness of the evaluation process induced by the decision procedure. Joint evaluation (JE) occurs when the options are presented simultaneously, and they are easily comparable; separate evaluation (SE) corresponds to opposite situations in which options are presented and evaluated one after the other. Hsee *et al*. show that a shift between JE and SE mode is sufficient to induce PR, holding constant the worth scale. They explain the effect by resorting to a notion of *evaluability*: some attributes are easier to evaluate in isolation, while others can be fully apprehended and appreciated only by comparing the options. Easily evaluable attributes are likely to play a prominent role in the SE mode, while attributes with lower evaluability will enter the evaluation process only in the comparative JE mode. According to Hsee *et al*., the difference in attributes considered in the JE and SE mode explains the emergence of the PR effect.

While Hsee *et al*. carefully restrict their explanation to the case in which the worth scale is held constant, it is tempting to stretch their argument and suggest that in many cases the effect might persist over different worth scales. The goal of this paper is exactly to explore under which interactions between response methods and worth scale the PR phenomenon is amplified or reduced. We ran two experiments where subjects were asked to evaluate attractiveness and minimum selling price both using the same response method: JE or SE.

The objects of evaluation in this study are lottery gambles, the most classic object of the PR literature. However, gambles are quite complex informational objects: they contain information on probabilities and outcomes, and both are needed to compute (even approximately) expected values and variance (or other risk indicators). This might make gambles ill suited for a study motivated by the evaluability theory: too many evaluability factors might affect the impact of different worth scales on different evaluation methods. Nevertheless, Ganzach (1996) has recently introduced simpler gambles (LVG vs. HVG) with equiprobable outcomes and nearly equal expected values (for simplicity, we will refer to such gambles with the acronym EOEEV-gambles, EOEEV for ***e***quiprobable ***o***utcomes and ***e***qual ***e***xpected ***v***alues). He has shown that the PR phenomenon holds also with EOEEV-gambles: although subjects choose LVG (in which all outcomes are moderate), they put a higher price on HVG (with very high as well as very low outcomes). These gambles considerably reduce cognitive difficulties connected to probabilities, leaving all variance to be determined by the values of different outcomes and making the calculation of expected values a rather simple arithmetic task. We will use similar gambles in our experimental design.

Then, applied to EOEEV-gambles, the evaluability hypothesis can be stated in the following way. EOEEV-gambles have only outcomes as sources of differentiation between single bets, since the role of probability is eliminated by the use of equiprobable gambles. If one restricts his attention only to outcomes, since EOEEV-gambles also have equal expected values, the main source of differentiation is the value of each outcome, the most salient outcomes, and the variance (or other variability measures) of outcomes in each gamble. We expect that under SE, variance is the less evaluable attribute of an EOEEV-gamble; in this case, salient values should play a key role in the evaluation process, and former experiments suggest that anchoring should occur on the highest monetary outcomes of each gamble (cf. Lichtenstein and Slovic, 1971; Schkade and Johnson, 1989; Ganzach, 1996). Consequently, HVG, which have the highest monetary outcomes, should have a comparative advantage under SE. Under JE, however, it shouldn't be too hard to compare the variance of gambles. Risk considerations should play a larger role in the evaluation process, and risk-averse behavior should be expected to prevail, at least as far as gambles are defined in the domain of gains (Kahneman and Tversky, 1979).

Furthermore, we expect that the worth scales should have a minor effect on evaluability: whether the scale is attractiveness or minimum selling price, variance is still the hardest attribute to evaluate, and the JE/SE distinction should be dominant over worth scale differences.

Consequently, all four response types could be obtained as the modal ones, according to how JE and SE mode are combined. In particular, we hypothesize that in the JE mode, the variance of equiprobable gambles will be taken into account, inducing more risk-averse behaviors, while in the SE mode, anchoring on the highest outcomes would become prominent, inducing more risk-seeking behaviors. We expect this effect to be robust, i.e., to persist under variation in the classic worth scales. Exhibit 2 summarizes our predictions.

[insert Exhibit 2 around here]

However, the evaluability theory emphasizes the relative difficulty in evaluating information on different attributes as a source of PR. This should be reflected in the processes through which subjects search and compare information on gambles. A self-paced display time paradigm (SDTP) allows to trace these processes. Subjects are presented with slots of covered information on a computer screen, and have to actively uncover it by passing the mouse on the slots or by pressing keys. This way, it is possible to track how much time subjects spend looking at each piece of stimuli, and what sequence of information slots they go through (cf. Caverni, 1987; Schkade and Johnson, 1989; Payne *et al.*, 1992). Thus, together with the fact that there is a lack of observable in the PR literature, it was legitimate to include SDTP in our experimental design.

EXPERIMENT 1: JOINT EVALUATION VERSUS SEPARATE EVALUATION WITH ATTRACTIVENESS AND MINIMUM SELLING PRICE WORTH SCALES

The aims of Experiment 1 was to verify that (1) the PR phenomenon still occurred with EOEEV-gambles when making use of the classical elicitation procedures (control group), (2) the classic PR was reduced when a same evaluation mode was used for both attractiveness and minimum selling price worth scales (JE and SE groups), and (3) both anchoring and evaluability had an effect on the allocation of attention (control, JE and SE groups).

In each group, each subject evaluated the stimuli in both the attractiveness and the minimum selling price conditions. Within each group, half of the subjects dealt with the attractiveness condition first, while the other half dealt with the minimum selling price condition first.

Method

Participants

Ninety-six students in psychology participated, 32 in each group: in the control and JE groups, 20 subjects were from the University of Quebec in Montreal and 12 from the Aix-Marseille I University (AMU); in the SE group all subjects were from the AMU.

Stimuli

The stimuli were 16 gambles with equiprobable outcomes and equal expected values (EOEEV-gambles). Each gamble involved four possible outcomes ranging from the lower to the higher. The outcomes were either in French francs (FF) or in Canadian dollars ($C) according to the subjects' nationality: the lowest outcome was 3FF (1$C), the highest was 368FF (109$C). The lottery consisted in selecting one and only one of the four possible outcomes of each gamble. Thus, for each gamble, subjects were sure to win one of the four outcomes and the probability to win one of each of these outcomes was 0.25.

For example, if the first outcome was selected, subjects could only win 7 FF with a gamble involving the outcomes 7, 31, 56 or 106 FF (HVG), while they could win 34 FF with a gamble involving the outcomes 34, 46, 53 or 67 FF (LVG). If the fourth outcome was selected, subjects could win 106 FF with the HVG while they could only win 67 FF with the LVG.

Each pair of these gambles included a LVG and a HVG with equal expected values (50FF ($C15) or 150FF ($C45)). The range of each HVG was always three times larger than the range of LVG, and the number of outcomes dominating the corresponding ones in the other gamble of the same pair was manipulated: for one pair, the LVG had three outcomes higher than the HVG; for another pair, the HVG had three outcomes higher than the LVG; and for two others, LVG and HVG had an equal number of dominating outcomes (cf. Exhibit 3).

[insert Exhibit 3 around here]

Material

The experiment was run on a PC and the SDTP was carried out by a C++ program imposing a sequential information processing. Gambles were presented masked. In order to look at each outcome of each gamble, subjects had to press the appropriate colored keys on the keyboard, one key corresponding to one and only one outcome. Each outcome remained visible until the subjects stopped to press the corresponding key. It was impossible to look at two outcomes at the same time but subjects could look at them as many times and for as long as they wanted. Another keys were used to give the response. When subjects made visible an outcome on the screen, the program recorded which outcome it was and the time spent on it.

Procedure

Exhibit 4 shows the experimental design of Experiment 1. In JE mode (binary response scale), subjects were asked to indicate, for each pair, either the gamble they would prefer in order to participate in the lottery (attractiveness condition of control and JE groups) or the gamble for which they would attribute the highest minimum selling price in order to give up their participation in the lottery (pricing condition of the JE group). When they did their choice, subjects had to press the appropriate key in order to indicate the chosen gamble. Then a new pair of gambles (first masked) appeared. In SE mode (numerical response scale), subjects were asked to indicate, for each gamble, either their degree of attraction (this procedure is known as the "rating task" in literature) on a scale between 0 and 10 (“0” for “Not attractive” and “10” for “Very attractive”) in order to participate in the lottery (attractiveness condition of the SE group) or the minimum selling price they would require in order to give up their participation in the lottery (pricing condition of control and SE groups). When they did their evaluation, subjects had to press the appropriate key in order to indicate the value they had estimated. Then a new gamble (first masked) appeared.

The control group was used in order to verify whether the PR phenomenon would still hold with our subjects’ population and with our experimental material and procedure.

[insert Exhibit 4 around here]

Results

As it is usual in the PR literature, we considered as "*preferred gambles*" those that were chosen when presented pair by pair, and those that received the highest value when presented separately. We ran two kinds of analysis. One was on the *given preferences*, in which the dependant variable was the percentages of the HVG preferred by each subject ($\%_{HVG} = n_{HVG} / n_{LVG} + n_{HVG}$). The other was on the *information taking*, i.e. the visualization time of each outcome for each gamble.

Preferences

Exhibit 5 summarizes the results. In the control group, the PR phenomenon is neatly reproduced. Under the attractiveness condition the LVG are preferred, while under the pricing condition preferences are reversed. The percentages (64 vs. 34) of HVG preferred with the two worth scales are significantly different ($F(1,31)=21.52$; $p<.0001$)[2]. Furthermore, differences between the percentage (41) of the classic PR and the percentage (11) of the unexpected PR (cf. Exhibit 2) are significant as well ($F(1,31)=21.52$; $p<.0001$)[3]. But, JE and SE groups show that consistent patterns of response of both the risk-averse and the risk-seeking type can be obtained when the response method is held constant through the evaluation mode: there are no significant differences between the rates of HVG

[2] All F-tests were performed at an alpha level of 0.01.

[3] According to our design, these two F-tests are the same.

preferred nor between those of PR in JE ($F(1,31)=2.79$; $p<.1050$) and SE ($F(1,31)=.07$; $p<.7996$) groups.

[insert Exhibit 5 around here]

Between-subject comparisons of our different groups show a remarkable overall robustness of each experimental outcome. A comparison between control and JE groups shows that while there are no significant differences between the percentages (36 vs. 34) of HVG preferred in the JE/attractiveness combinations, the percentage (64) in the SE/minimum selling price combination is significantly different from the percentage (46) given in the JE/minimum selling price combination ($F(1,62)=5.49$; $p<.0223$). The comparison between control and SE group shows a complementary pattern: while there are no significant differences between the percentages (64 vs. 62) of HVG preferred in the SE/minimum selling price combinations, the percentage (34) in the JE/attractiveness combination is significantly different from the percentage (63) given in the SE/attractiveness combination ($F(1,62)=20.71$; $p<.00001$).

Further insight can be gained by comparing the rate of classic PR responses between the 3 groups: the percentages of this inconsistency are significantly lower in JE ($F(1,62)=5.82$; $p<.0188$) and SE ($F(1,62)=9.19$; $p<.0035$) groups than in the control group. Moreover, while the modal pattern of response in the control group is the classic PR one, in the JE group the modal pattern of response is the consistent risk-averse one, and in the SE group it is the consistent risk-seeking one. This demonstrates that an effective elimination[4] of the PR phenomenon has been obtained.

Information taking

While the above results seem to provide overall support to our hypotheses, process observations provided by the use of the SDTP supply relevant insights in the actual patterns of attention related to subjects' responses. Former process analysis of PR (Schkade and Johnson, 1989) have shown that the hypothesis relating different information processing patterns to the worth scales is supported by differences in the average response times. The minimum selling price condition exhibits systematically higher response times than attractiveness condition (choice or rating). Our experiment has obtained similar findings in terms of visualization times. In all three groups (cf. Exhibit 6) the minimum selling price condition implies visualization times that are approximately 50% higher than the attractiveness condition ($F(1,228)=64.99$; $p<.00001$ for the control group, $F(1,255)=41.30$; $p<.00001$ for the JE group and $F(1,180)=76.36$; $p<.00001$ for the SE group).

[insert Exhibit 6 around here]

But how can process data be related to our main results? In fact, the most interesting observations arise at a finer grain of analysis. The subjects do not look with the same attention at the outcomes of gambles whether they preferred LVG or HVG. While the

[4] That means some reversal still occurs but they are no longer systematic classic PR.

pattern of visualization time of outcomes shows a quite flat distribution of attentional efforts when LVG are preferred, it exhibits a sharply inclined curve, monotonically increasing with outcome values when the HVG are preferred (cf. Exhibit 7: the interaction between the preference and the outcomes visualization time is significant ($F(3,3951)=19.50$; $p<.0000$)). This effect is observed whatever the worth scale and the evaluation mode may be[5].

[insert Exhibit 7 around here]

Discussion

These results confirm the existence of the PR phenomenon even when highly simplified gambles such as EOEEV-gambles are the object of evaluation. Furthermore, when both worth scales are combined with a JE mode, the modal response is consistently risk-averse, while when both worth scales are combined with a SE mode, modal response behaviors shift towards risk-seeking consistency.

Our results suggest that when all dimensions of the stimuli (i.e. outcome variance or some approximation of it and the value of each outcome) are taken into account by subjects, their behavior tends to be risk-averse. On the opposite, when only one dimension of the stimuli (i.e. the high outcomes) is taken into account by subjects, their behavior tends

[5] Results did not differ significantly between the three groups; then, the analysis has been collapsed over this factor.

to be risk-seeking. This nicely fits into the anchoring explanation for the classic PR phenomenon, originally set forth by Slovic and Lichtenstein (1968), even if rather than suggesting that only the highest value is considered, our data suggest a smoother process of attraction towards high value outcomes. The consequence is that when the response method involves comparison (binary response scale) between the stimuli (and thus the outcome variance is easier to integrate, having a greater evaluability), the low-variance options are preferred. Conversely, the high-variance options are preferred when the response method does not allow comparison between stimuli. As a result, the PR phenomenon, observed in our control group, disappears in our experimental groups, the modal response shifting either towards consistent risk-averse behavior (JE group) or towards consistent risk-seeking behavior (SE group). This is consistent with the evaluability theory.

Finally, we observe that the average time spent on visualizing each gamble is strikingly similar for JE and SE groups in the attractiveness condition. This seems to support, at least at this level of aggregation, that it is justified to assign both classic choice on a unique attractiveness worth scale type, as Goldstein and Einhorn (1987) already did. We emphasize that we are not claiming that choice and rating are the same kind of activity (see Schkade and Johnson, 1989; Mellers *et al.*, 1992, for some strong negative evidence), but we suggest instead that it is legitimate to claim that they resort to similar worth scales, implying similar attentional effort.

EXPERIMENT 2: EXPLICIT COMPARISON VERSUS NO EXPLICIT COMPARISON IN JOINT EVALUATION

Our second experiment has been designed to answer to a possible objection to Experiment 1, while refining the evaluability hypothesis. A possible weakness in the design of Experiment 1 concerns the evaluation of the minimum selling price in the JE group. In such condition, subjects were asked to indicate, for each pair of gambles, the gamble for which they would ask the higher minimum selling price. One might object that such task does not necessarily require that subjects actually estimate the minimum selling price of both gambles – they might just answer by indicating their preferred one. In such case, the task would be much closer to the attractiveness condition than to the pricing one. The consistent risk-averse behavior observed in the JE group might thus reflect of a design fault. Thus, we designed a variation of the original JE group in which subjects, besides evaluating attractiveness as in Experiment 1, would have to indicate, for each pair of gambles, the gamble for which they would ask the higher minimum selling price and they would have to write the minimum selling price for each gamble. Our hypothesis was of course that the behavior observed in this case would be very similar to the one observed in the JE group of Experiment 1. We labeled this group the "JE-Explicit comparison group".

At the same time, we wanted to refine our understanding of the evaluability hypothesis; in particular, we were interested in assessing how much the evaluability effect is due to an evaluation mode effect (JE vs. SE) or to the comparative nature of the task. We thus presented to a second group of subjects, maintaining the JE mode, two tasks. One was just the attractiveness condition of Experiment 1. In the second task, subjects were asked only to indicate, for each pair of gambles, a minimum selling price for each gamble, without any explicit comparison between gambles being asked. We labeled this group the “JE-No explicit comparison group”. Our hypothesis was that in absence of an explicitly comparative task, the effects of the JE mode would be significantly weakened.

Method

The stimuli (only in FF) and the material were the same as in Experiment 1. Each subject evaluated the stimuli in both attractiveness and minimum selling price conditions.

Participants

Sixty (28 in the JE-Explicit comparison group and 32 in the JE-No explicit comparison group) students in psychology (baccalaureate) of the AMU participated. Within each group, half of the subjects dealt with the attractiveness condition first, while the other half dealt with the minimum selling price condition first.

Procedure

Exhibit 8 shows the experimental design of Experiment 2. The attractiveness condition was the same in JE-Explicit and JE-No explicit comparison groups, as in control and JE groups of Experiment 1. In the pricing condition of the JE-Explicit comparison group (binary and numerical response scale), subjects were asked to indicate, for each gamble, first the gamble for which they would attribute the highest minimum selling price in order to give up their participation in the lottery, and, after only, the price of each gamble. Then, a new pair of gambles (first masked) appeared. In the pricing condition of the JE-No explicit comparison group (numerical response scale), subjects were asked to indicate, for each gamble, the minimum selling price they would require in order to give up their participation in the lottery.

[insert Exhibit 8 around here]

Results

Preferences

Exhibit 9 summarizes the results. On the one hand, they show that, as expected, the JE-Explicit comparison group behaves just like the corresponding JE group of Experiment 1, exhibiting consistent risk-averse behavior with the two worth scales. There are no significant differences between the rate of HVG preferred, nor between those of PR ($F(1,27)=2.15$; $p<.1538$). On the other hand, in the JE-No explicit comparison group, the PR phenomenon re-appears as the pricing behavior becomes risk-seeking when the comparative aspects of the task are made only implicit. The percentages (63 vs. 33) of HVG preferred with the two worth scales and those (42 vs. 12) of PR are significantly different ($F(1,31)=17.50$; $p<.0002$).

[insert Exhibit 9 around here]

Between-subject comparisons of the different groups of the Experiment 2 show here again a great overall robustness of each observation. Thus, between-subjects comparisons of JE-No explicit comparison and JE-Explicit comparison groups shows that while there are no significant differences between the percentages (35 vs. 33) of HVG preferred in the attractiveness condition, the percentage (63) in the JE-No explicit comparison/minimum selling price combination is significantly different from the percentage (43) given in the JE-explicit comparison/minimum selling price combination ($F(1,58)=5.95$; $p<.0178$), and also from the percentage (46) observed in the JE group ($F(1,62)=4.32$; $p<.0419$). However, there are no significant differences between the percentages of HVG preferred of both JE-No explicit comparison and control groups, and JE-Explicit comparison and JE groups. The comparison of control and JE-Explicit comparison groups shows a complementary pattern: while there are no significant differences between the percentages (35 vs. 34) of HVG preferred in the attractiveness condition, the percentage (64) in the minimum selling price condition of the control group is significantly different from the percentage (43) given in the JE-Explicit comparison/minimum selling price combination ($F(1,58)=7.56$; $p<.0080$).

By comparing the rate of classic PR responses between these two groups, we observe the same kind of analysis: the percentages of this inconsistency is significantly lower (42 vs. 25) in the JE-Explicit comparison group than in the JE-No explicit comparison group ($F(1,58)=6.97$; $p<.0107$), and also (41 vs. 25) than the control group ($F(1,58)=6.06$; $p<.0168$). However, there are no significant differences between the classic PR percentages of JE-Explicit comparison and JE groups or the SE group, and those of JE-

No Explicit comparison and control group, demonstrating that an effective elimination[6] of the PR phenomenon has been obtained in this experiment too.

Information taking

We obtain the same effect as in Experiment 1. First, concerning the global visualization times: in the two groups ($F(1,223)=137.53$; $p<.0000$ for the JE-Explicit comparison group and $F(1,222)=102.50$; $p<.0000$ for the JE-No explicit comparison group), the minimum selling price condition implies visualization times approximately 50% higher significantly than the attractiveness condition (cf. Exhibit 10).

[insert Exhibit 10 around here]

Second, concerning the time spent on each outcome: when subjects prefer the LVG, each outcome is visualized with similar attention effort, while we observe an increasing focalization on the highest values when subjects prefer HVG (cf. Exhibit 11: the interaction between the preference and the outcomes visualization time is significant ($F(3,2625)=16.36$; $p<.0000$)).

[insert Exhibit 11 around here]

[6] Cf. note 4.

GENERAL DISCUSSION

By showing both the presence of the PR phenomenon (control and JE-No explicit comparison groups) with EOEEV-gambles (confirming the usefulness of these stimuli) and its elimination (JE, SE and JE-Explicit comparison groups), our results support the hypothesis that the shift in the response method can be the major cause of such a phenomenon, even across different worth scales, suggesting the minor role of this latter factor. Furthermore, they show that different individual preferences, affected by response method, also have an impact on the distribution of attention, as revealed by visualization times. This confirms the usefulness of the SDTP, even beyond the pioneering results already obtained by Schkade and Johnson (1989), who had found weaker relations between process tracing measures and PR; we suggest that the simpler gamble type used in the present experiments might be responsible for this neater evidence.

These observations confirm the effect of the evaluation mode (JE vs. SE) and then are consistent with the evaluability theory (Hsee *et al.*, 1999): the way subjects evaluate is influenced by the given information and also by how they have to combine it. But, they also show that anchoring processes actually occur, and suggest that their emergence is related to the difficulty of evaluating variance in the SE mode: we have seen that steep attention curves are frequent when HVG are preferred, i.e. in the SE mode, independently from the worth scale. But if anchoring plays an important role in the emergence of the phenomenon, it is not its ultimate cause; as we have seen in the SE group, when two different worth scales are used under a same SE mode, anchoring occurs with both worth scales and the modal response type is consequently consistently risk-seeking. Then, the PR phenomenon is eliminated.

However, Experiment 2 suggests that, in the JE mode, the mere joint availability of information on pairs of options may not suffice to trigger the use of less evaluable attributes; subjects may still ignore comparable information. Instead, actual comparison needs to be induced by the JE task. This suggests that requirements on actual comparability of pairs of options may be more stringent than in Hsee *et al*. (1999), emphasizing that it is the actual presence/absence of a comparing process, rather than its possibility, which is responsible for the classic PR.

Goldstein and Einhorn (1987) distinguish three stages of response processes: encoding, evaluation and expression, and attribute to the last stage the most critical role in determining the classic PR. In contrast, the correspondence of individual preferences and attentional patterns observed in our experiments suggests that much happens in the early stages. Once more, we suggest a cautious interpretation of our results. In particular, when more complex stimuli are used, the expression stage might be more relevant and even have significant feedback on encoding and evaluation stages. This might explain some significant divergences between our results (and also those of Hsee *et al*., 1999) and the original results by Goldstein and Einhorn (1987). In particular, even under a same response method Goldstein and Einhorn find patterns of response contrasting with those predicted and obtained in Hsee *et al*. and our experiments.

CONCLUSION

Evaluability and anchoring hypothesizes seem to be good explanations for the emergence of the classic PR with EOEEV-gambles. The classical conditions used to elicit preferences reveal the PR phenomenon because different evaluation modes are proposed to subjects in each condition. So, when we asked them to compare the gambles (JE mode), they can take into account all the dimensions of the stimuli, but they can't when they evaluate gambles one by one (SE mode) and then focus on the high outcomes by an anchoring process. One important consequence is that different worth scales (attractiveness vs. minimum selling price) have no influence on the subjects preferences here, even if visualization times reveal different processes for each worth scale.

A final reflection about our findings lead us to amplify a doubt already implicit in this notion of evaluability: is the classic PR a phenomenon genuinely concerning the construction of individual preferences, or does it reflect (at least in many cases) the use of different information in the response processing? The latter hypothesis is reinforced by the fact that the PR phenomenon tends to become less frequent under appropriate manipulations of response methods. If this is the case, the PR phenomenon may turn out to be an epiphenomenona of underlying information processing rather than a true phenomenon of cognitive inconsistency.

ACKNOWLEDGMENTS

We thank Nicolas Lipari for his C++ program conception and Héloïse Joly for his very helpful reading. The Cognitive Sceince Lab of Rovereto and the Cognitive Psychology Laboratory of Aix-Marseille provided warm hospitality to two of the authors, making our joint effort possible.

REFERENCES

Camerer C. 1995. Individual decision making. In *The Handbook of Experimental Economics,* Kagel JH, Roth AE (Eds). Princeton: Princeton University Press.

Caverni J-P. 1987. Self-paced display time for process-tracing in assessment of acquired knowledge. *European Bulletin of Cognitive Psychology* **7**: 633-651.

Caverni J-P. 1996. How to better understand the processes underlying the so-called « preference reversal phenomenon » if there is any reversal phenomenon? *Journal of Behavioral Decision Making* **9**: 111.

Ganzach Y. 1996. Preference reversals in equal-probability gambles: a case for anchoring and adjustment. *Journal of Behavioral Decision Making* **9**: 95-109.

Goldstein WM, Einhorn HJ. 1987. Expression theory and the preference reversal phenomena. *Psychological Review* **94**: 236-254.

Grether DM, Plott CR. 1979. Economic theory and the preference reversal phenomenon. *American Economic Review* **69**: 623-638.

Hsee CK, Loewenstein GF, Blount S, Bazerman MH. 1999. Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological Bulletin* **125**: 576-590.

Kahneman D, Tversky A. 1979. Prospect theory: an analysis of decision under risk. *Econometrica* **47**: 263-291.

Lichtenstein S, Slovic P. 1971. Reversal of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology* **89**: 46-55.

Mellers BA, Chang S-J, Birnbaum MH, Ordóñez LD. 1992. Preferences, prices and ratings in risky decision making. *Journal of Experimental Psychology: Human Perception and Performance* **18**: 347-361.

Ordóñez LD, Mellers BA, Chang S-J, Roberts J. 1995. Are preference reversals reduced when made explicit? *Journal of Behavioral Decision Making* **8**: 265-277.

Payne JW, Bettman JR, Johnson EJ. 1992. Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology* **43**: 87-131.

Pommerehne W, Schneider F, Zweifel P. 1982. Economic theory of choice and the preference reversal phenomenon: A reexamination. *American Economic Review* **72**: 569-574.

Reilly RJ. 1982. Preference reversal: further evidence and some suggested modifications in experimental design. *American Economic Review* **72**: 576-584.

Schkade DA, Johnson EJ. 1989. Cognitive processes in preference reversals. *Organizational Behavior and Human Decision Processes* **44**: 203-231.

Slovic P, Lichtenstein S. 1968. Relative importance of probabilities and payoffs in risk taking. *Journal of Experimental Psychology Monographs* **78**: 165-182.

Slovic P, Lichtenstein S. 1983. Preference reversals: A broader perspective. *American Economic Review* **73**: 596-605.

Tversky A, Slovic P, Kahneman D. 1990. The causes of preference reversal. *American Economic Review* **80**: 204-217.

Authors' biographies

Christophe Schmeltzer is a Psychology Ph.D. student at the Aix-Marseille I University. His research interests are primarily in the area of judgment and decision making with a special interest in judgment bias.

Jean-Paul Caverni is a Professor in Cognitive/Experimental Psychology. He is the head of a research-team on the study of the inferential processes in reasoning and decision making. He currently works on hypothetico-deductive and probabilistic reasoning.

Massimo Warglien is a Professor of Information and Decision Making at the Ca' Foscari University of Venezia, and a Research Director at the Cognitive Science Lab of Rovereto.

Authors' addresses

Christophe Schmeltzer and **Jean-Paul Caverni,** Université d'Aix-Marseille I, Laboratoire de Psychologie Cognitive, Case 66, 3 Place Victor Hugo, F-13331 Marseille Cedex 3, France.

Massimo Warglien, Università Ca'Foscari di Venezia, Dipartimento di Economia e Direzione Aziendale, Dorsoduro 1075, 30123 Venezia, Italia.

EXHIBITS

Exhibit 1. Possible response types.
(LVG = low-variance gamble; HVG = high-variance gamble)

		Judgment	
		HVG	**LVG**
Choice	**HVG**	Consistent risk-seeking	Unexpected PR
	LVG	Classic PR	Consistent risk-averse

Exhibit 2. Our Preferences predictions.
(JE = joint evaluation; SE = separate evaluation)

		Minimum selling price	
		SE	**JE**
Attractiveness	**SE**	Consistent risk-seeking	Unexpected PR
	JE	Classic PR	Consistent risk-averse

Exhibit 3. Stimuli (FF): eight pairs of EOEEV-gambles with four possible outcomes, their expected value (EV), their range and their dominated outcomes (> or <).

Pair n°	Gamble type	Outcome 1	Outcome 2	Outcome 3	Outcome 4	EV	Range HVG/3 = LVG
1	**HVG**	4	41 >	61 >	94 >	**50**	**90**
	LVG	< 37	39	57	67		**30**
2	**HVG**	7	31	56 >	106 >	**50**	**99**
	LVG	< 34	< 46	53	67		**33**
3	**HVG**	3	46 >	49	102 >	**50**	**99**
	LVG	< 31	43	< 62	64		**33**
4	**HVG**	6	32	39	123 >	**50**	**117**
	LVG	< 29	< 37	< 66	68		**39**
5	**HVG**	13	122 >	182 >	283 >	**150**	**270**
	LVG	< 112	117	169	202		**90**
6	**HVG**	22	92	167 >	319 >	**150**	**297**
	LVG	< 103	< 137	158	202		**99**
7	**HVG**	11	139 >	142	308 >	**150**	**297**
	LVG	< 92	131	< 186	191		**99**
8	**HVG**	17	96	119	368 >	**150**	**351**
	LVG	< 86	< 114	< 197	203		**117**

Exhibit 4. Experimental design of Experiment 1.
(b = binary response scale; n = numerical response scale)

		Group		
		Control group	**JE group**	**SE group**
Worth scale	**Attractiveness**	Pair by pair (b)	Pair by pair (b)	One by one (n)
	Minimum selling price	One by one (n)	Pair by pair (b)	One by one (n)

Exhibit 5. Worth scales and averages of percentages of HVG preferred (gray), of PR (bulk), and of the modal response (*), for each group of Experiment 1.

Control group

		Minimum selling price		
		HVG	**LVG**	
Attractiveness	**HVG**	23	**11**	34
	LVG	**41***	25	66
		64	36	100

JE group

		Minimum selling price		
		HVG	**LVG**	
Attractiveness	**HVG**	21	**15**	36
	LVG	**25**	39*	64
		46	54	100

SE group

		Minimum selling price		
		HVG	**LVG**	
Attractiveness	**HVG**	40*	**23**	63
	LVG	**22**	15	37
		62	38	100

Exhibit 6. Visualization time (ms) of each gamble for each group of Experiment 1.

		Group		
		Control group	**JE group**	**SE group**
Worth scale	**Attractiveness**	1157	1184	1281
	Minimum selling price	1905	1721	2312

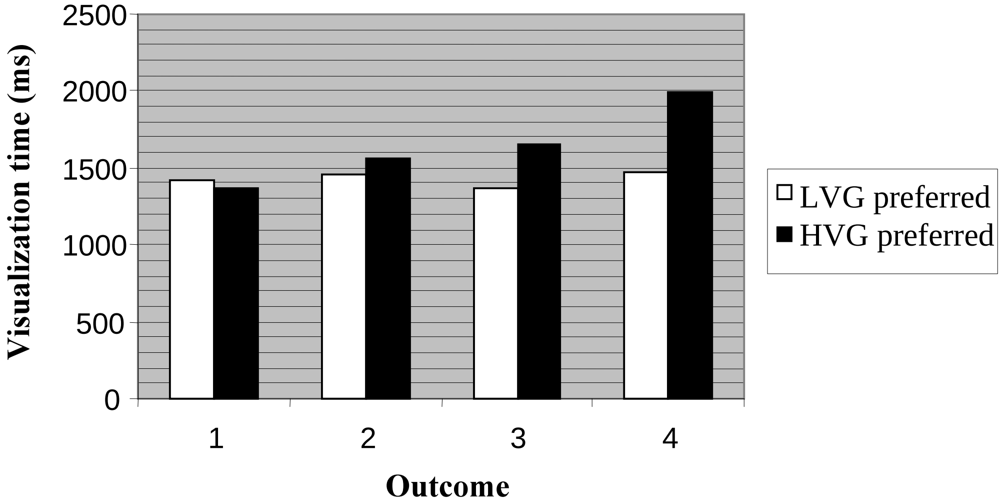

Exhibit 7. Graph interaction of visualization time of each outcome, between preference and outcome, for all groups of Experiment 1.

Exhibit 8. Experimental design of Experiment 2.
(b = binary response scale; n = numerical response scale)

		Group	
		JE-Explicit comparison group	**JE-No explicit comparison**
Worth scale	**Attractiveness**	Pair by pair (b)	Pair by pair (b)
	Minimum selling price	Pair by pair (b, n)	Pair by pair (n)

Exhibit 9. Worth scales and averages of percentages of HVG preferred (gray), of PR (bulk), and of the modal response (*), for each group of Experiment 2.

JE-Explicit comparison group

		Minimum selling price		
		HVG	**LVG**	
Attractiveness	**HVG**	18	**19**	35
	LVG	**21**	38*	65
		43	57	100

JE-No explicit comparison group

		Minimum selling price		
		HVG	**LVG**	
Attractiveness	**HVG**	21	**12**	33
	LVG	**42***	25	67
		63	37	100

Exhibit 10. Visualization time (ms) of each gamble for each group of Experiment 2.

		Group	
		JE-Explicit comparison group	JE-No explicit comparison group
Worth scale	**Attractiveness**	1173	1253
	Minimum selling price	2577	3039

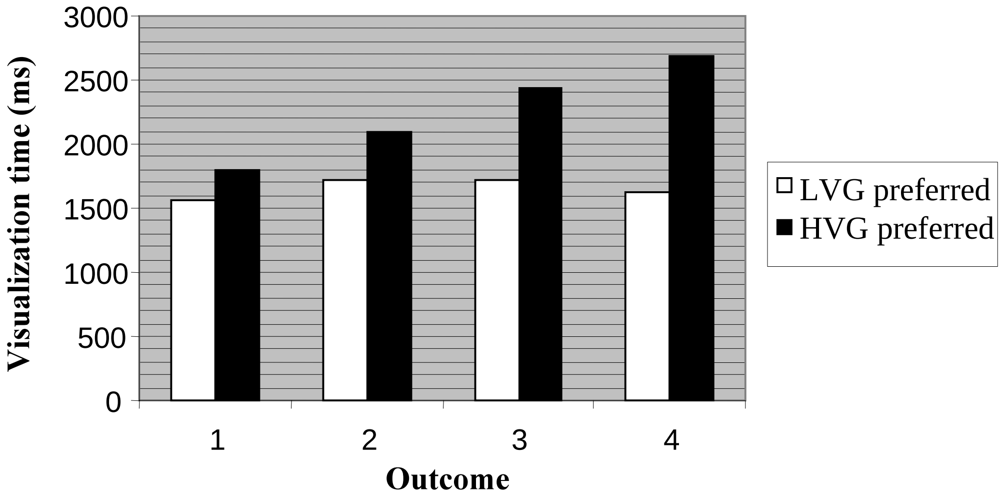

Exhibit 11. Graph interaction of visualization time of each outcome, between preference and outcome, for all groups of Experiment 2.